EXHIBIT 6

Int Urogynecol J (2015) 26:1155–1159
DOI 10.1007/s00192-015-2661-5

ORIGINAL ARTICLE

Missing data frequency and correlates in two randomized surgical trials for urinary incontinence in women

Linda Brubaker · Heather J. Litman · Hae-Young Kim · Philippe Zimmern · Keisha Dyer · John W. Kusek · Holly E. Richter · Anne Stoddard · for the Urinary Incontinence Treatment Network

Received: 23 October 2014 / Accepted: 6 February 2015 / Published online: 24 March 2015

Abstract
Introduction and hypothesis Missing data is frequently observed in clinical trials; high rates of missing data may jeopardize trial outcome validity.
Purpose We determined the rates of missing data over time, by type of data collected and compared demographic and clinical factors associated with missing data among women who participated in two large randomized clinical trials of surgery for stress urinary incontinence, the Stress Incontinence Surgical Treatment Efficacy Trial (SISTEr) and the Trial of Midurethral Sling (TOMUS).
Methods The proportions of subjects who attended and missed each follow-up visit were calculated. The chi-squared test, Fisher's exact test and *t* test were used to compare women with and without missing data, as well as the completeness of the data for each component of the composite primary outcome.
Results Data completeness for the primary outcome computation in the TOMUS trial (62.3 %) was nearly double that in the SISTEr trial (35.7 %). The follow-up visit attendance rate decreased over time. A higher proportion of subjects attended all follow-up visits in the TOMUS trial and overall there were fewer missing data for the period that included the primary outcome assessment at 12 months. The highest levels of complete data for the composite outcome variables were for the symptoms questionnaire (SISTEr 100 %, TOMUS 99.8 %) and the urinary stress test (SISTEr 96.1 %, TOMUS 96.7 %). In both studies, the pad test was associated with the lowest levels of complete data (SISTEr 85.1 %, TOMUS 88.3 %) and approximately one in ten subjects had incomplete voiding diaries at the time of primary outcome assessment. Generally, in both studies, a higher proportion of younger subjects had missing data. This analysis lacked a patient perspective as to the reasons for missing data that could have provided additional information on subject burden, motivations for adherence and study design. In addition, we were unable to compare the effects of the different primary outcome assessment time-points in an identically designed trial.
Conclusions Missing visits and data increased with time. Questionnaire data and physical outcome data (urinary stress test) that could be assessed during a visit were least prone to missing data, whereas data for variables that required subject effort while away from the research team (pad test, voiding diary) were more likely to be missing. Older subjects were more likely to provide complete data.

Keywords Missing data · Urinary incontinence · Clinical outcomes · Surgical trials · Primary outcome measures

Clinicaltrials.gov identifiers TOMUS: NCT00325039; SISTEr: NCT00064662

L. Brubaker (✉)
Departments of Ob/Gyn and Urology, Stritch School of Medicine, Loyola University of Chicago, 2160 South First Avenue, Maywood, IL 60153, USA
e-mail: Lbrubaker@lumc.edu

H. J. Litman
Boston Children's Hospital, Boston, MA, USA

H.-Y. Kim · A. Stoddard
New England Research Institutes, Watertown, MA, USA

P. Zimmern
University of Texas Southwestern, Dallas, TX, USA

K. Dyer
University of California, San Diego, CA, USA

J. W. Kusek
National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MA, USA

H. E. Richter
University of Alabama at Birmingham, Birmingham, AL, USA

Introduction

High-quality clinical trials are characterized by sound study design, consistent collection of complete information, and appropriate statistical analyses. Missing data in clinical trials has received increasing attention recently as it presents distinct challenges for data analysis and interpretation [1, 2]. Strategies to reduce missing data should be considered prior to trial initiation; ongoing monitoring of missing data during recruitment is also recommended to proactively reduce the amount of missing data [1]. Despite the common occurrence of missing data in clinical trials there have been few detailed reports on the factors that may influence the amount of missing data. Of particular interest is missing outcome data as study power may be diminished and statistical methods to address this problem are inadequate. The aim of this study was to determine the frequency of missing primary outcome data, the reasons why the data were not obtained, and study subject factors associated with missing data in two large randomized clinical trials of surgery for stress urinary incontinence, the Stress Incontinence Surgical Treatment Efficacy Trial (SISTEr), comparing outcomes between the Burch colposuspension and the pubovaginal sling, and the Trial of Midurethral Sling (TOMUS) study, comparing outcomes between the retropubic and transobturator midurethral sling approaches.

Methods

The design and findings of the SISTEr and TOMUS studies have been published previously [3–6]. All study subjects provided written informed consent and the institutional review board at each participating site approved the study protocols. The primary outcomes were assessed in the SISTEr and TOMUS studies at 24 and 12 months, respectively, after surgery/randomization. The primary outcome in each study was a composite measure including the Medical, Epidemiologic and Social Aspects of Aging Project (MESA) questionnaire to quantify the self-reported components of stress incontinence [7, 8], a self-completed 7-day voiding diary, urinary pad test and urinary stress test. In the urinary pad test urine loss is quantified by measuring the weight of a perineal pad worn by the subject during a standardized set of activities over a set period of time. We also collected data from the Incontinence Impact Questionnaire (IIQ) to evaluate the effect of incontinence on quality of life [7].

We determined the proportion of women who completed each study visit and the reasons for missing visit data (i.e., missed visit, subject withdrawal from trial, and refusal to provide data within an otherwise completed visit), as well as the pattern (timing) of missing data. If a woman was known to have met the definition of failure prior to the primary outcome visit, then her outcome data were complete for the time-to-event analysis even if she missed the primary outcome visit. A woman who did not meet the definition of failure prior to that visit and who did not attend the visit was censored at the last visit at which her outcome status was known. Retreatment was not a reason for withdrawal. Deaths were included in subjects lost to follow-up. The chi-squared test, Fisher's exact test and *t* test were used to compare the two groups (no missing vs. missing), as appropriate. In addition, we assessed the completeness for each component of the composite primary outcome variable. Analysis was performed using SAS version 9.3 (SAS Institute, Cary, NC). A 5 % two-sided significance level was used for all statistical testing.

Results

Table 1 shows the follow-up visit attendance rates and reasons for missing a visit, demonstrating that the rate of visit attendance decreased over time. A higher proportion of subjects attended all follow-up visits in the TOMUS study (Table 2) and, overall, there were fewer missing data for the period that included collection of the primary outcome at 12 months.

At the time of primary outcome assessment, a similar proportion of subjects completed the quality of life questionnaires (SISTEr 95.5 %, TOMUS 97.8 %). The highest proportions of complete data for the composite outcome variables were for the MESA questionnaire (SISTEr 100 %, TOMUS 99.8 %) and the urinary stress test (SISTEr 96.1 %, TOMUS 96.7 %). In both studies, the urinary pad test was associated with the lowest proportions of complete data (SISTEr 85.1 %, TOMUS 88.3 %) and approximately one in ten subjects had incomplete voiding diaries (Table 3).

Table 1 Visit attendance rates, and reasons for missing a visit

	SISTEr study (*n*=655 randomized)				TOMUS study (*n*=597 randomized)		
	6 months	12 months	18 months	24 months	6 months	12 months	24 months
Visit attended, *n* (%)	580 (88.5)	532 (81.2)	534 (81.5)	510 (77.9)	553 (92.6)	539 (90.3)	488 (81.7)
Withdrew/refused/lost to follow-up, *n* (%)	34 (5.2)	62 (9.5)	75 (11.5)	116 (17.7)	26 (4.4)	48 (8.0)	104 (17.4)
Visit missed, *n* (%)	41 (6.3)	61 (9.3)	46 (7.0)	29 (4.4)	18 (3.0)	10 (1.7)	5 (0.8)

Table 2 Patterns of visit attendance

	SISTEr study (*n*=655 randomized)	TOMUS study (*n*=597 randomized)
Primary outcome visit (months)	24	12
Completed all follow-up visits, *n* (%)	439 (67.0)	523 (87.6)
Completed primary outcome visit but missed one or more previous visits, *n* (%)	71 (10.8)	16 (2.7)
Withdrew, refused or lost to follow-up prior to primary outcome visit, *n* (%)	116 (17.7)	48 (8.0)
Missed primary outcome visit but not withdrawn, refused or lost to follow-up, *n* (%)	29 (4.4)	10 (1.7)

Table 4 shows the proportions of subjects who provided data on each component of the primary outcome at the time of assessment. The proportion of subjects who provided data on all four components in the TOMUS study (62.3 %) was nearly double that in the SISTEr study (35.7 %), without major differences in terms of surgical success. A higher proportion of subjects were censored prior to their primary outcome visit in the SISTEr study (20.6 %) than in the TOMUS study (6.0 %).

In both studies, a higher proportion of younger subjects had missing primary outcome data (Table 5). There were also study-specific associations: in the SISTEr study subjects with missing data were less likely to have had an adverse event and more likely to have had surgical retreatment of stress urinary incontinence, whereas in the TOMUS study subjects with missing data were less likely to have had a urinary incontinence episode at baseline.

Discussion

The studies analyzed in this report had a relatively high rate of missing data involving the primary outcome composite measure. Our analysis quantified and corroborated the widely held belief that missing data increases over time. This finding may assist other clinical investigators in balancing the desire for long-term outcomes with the reality of missing data. Good conduct of a clinical trial places a high value on timely collection of all planned data, although a priority is typically placed on collection of primary outcome data.

Table 3 Completeness of data among those who attended the primary outcome visit

	SISTEr (24 months) (*n*=510)	TOMUS (12 months) (*n*=539)
IIQ completed, *n* (%)	487 (95.5)	527 (97.8)
Primary outcome measures, *n* (%)		
MESA stress index	510 (100.0)	538 (99.8)
Voiding diary	450 (88.2)	490 (90.9)
Pad test	434 (85.1)	476 (88.3)
Stress test	490 (96.1)	521 (96.7)

It is possible that the lower follow-up visit attendance rate in the SISTEr study than in the TOMUS study was due to the longer duration of follow-up in the SISTEr study, as it is harder to sustain follow-up compliance over a longer time period. Nevertheless, even at the 12-month visit the follow-up visit attendance rate in the SISTEr study was slightly lower than in the TOMUS study. There was also likely an order effect, with the TOMUS study following the SISTEr study. The research teams were likely to have improved their skills in subject retention and their ability to obtain complete data. Another reason for the differences between the two studies

Table 4 Completeness of data at the primary outcome visit

	SISTEr study (*n*=655 randomized)	TOMUS study (*n*=597 randomized)
Primary outcome visit (months)	24	12
Complete data at primary outcome visit, *n* (%)[a]	234 (35.7)	372 (62.3)
Success	185 (79.1)	309 (83.1)
Failure	49 (20.9)	63 (16.9)
Censored prior to primary outcome visit, *n* (%)	135 (20.6)	36 (6.0)
Failed prior to primary outcome visit, *n* (%)	286 (43.7)	189 (31.7)

[a] The definitions of success and failure were different in the two studies. In the SISTEr study, overall success was defined as no self-reported symptoms of urinary incontinence, an increase of less than 15 g in pad weight during a 24-h pad test, no incontinence episodes recorded in a 3-day diary, a negative urinary stress test (no leakage noted on examination during cough and Valsalva maneuvers at a standardized bladder volume of 300 ml), and no retreatment for urinary incontinence (including behavioral, pharmacologic, and surgical therapies). The definition of success specific to stress incontinence was limited to no self-reported symptoms of stress incontinence, a negative stress test, and no retreatment for stress incontinence. The TOMUS study used a composite primary outcome assessed at 12 months after randomization that included treatment success according to objective criteria and treatment success according to subjective criteria. The objective criteria were a negative provocative stress test, a negative 24-h pad test, and no retreatment (behavioral, pharmacologic, or surgical) for stress incontinence. The subjective criteria were absence of self-reported symptoms of stress-type urinary incontinence, as assessed with the use of the Medical, Epidemiological and Social Aspects of Aging questionnaire, no leakage recorded in a 3-day voiding diary, and no retreatment for stress incontinence.

Table 5 Patient demographics at baseline and clinical characteristics at baseline and after surgery in relation to missing data in the two studies

Variable	SISTEr study			TOMUS study		
	No missing data at 24 months (n=510)	Missing data at 24 months (n=145)	p value[c]	No missing data at 12 months (n=539)	Missing data at 12 months (n=58)	p value[c]
Race/ethnicity, n (%)			0.06			0.44
Hispanic	54 (10.6)	18 (12.5)		63 (11.7)	8 (13.8)	
Non-Hispanic, white	386 (75.7)	94 (65.3)		430 (79.8)	43 (74.1)	
Non-Hispanic, black	29 (5.7)	15 (10.4)		16 (3.0)	1 (1.7)	
Non-Hispanic, other	41 (8.0)	17 (11.8)		30 (5.6)	6 (10.3)	
Marital status, n (%)			0.19			0.37
Married/living as married	356 (69.8)	93 (64.1)		375 (69.6)	37 (63.8)	
Not married or other	154 (30.2)	52 (35.9)		164 (30.4)	21 (36.2)	
Education, n (%)			0.09			0.25
Less than high school	40 (7.8)	14 (9.7)		28 (5.2)	7 (12.1)	
High school	134 (26.3)	37 (25.5)		137 (25.4)	12 (20.7)	
More than high school	194 (38.0)	68 (46.9)		195 (36.2)	22 (37.9)	
BA/BS	80 (15.7)	18 (12.4)		91 (16.9)	10 (17.2)	
Graduate/professional	62 (12.2)	8 (5.5)		88 (16.3)	7 (12.1)	
Concomitant surgery, n (%)			0.17			0.92
Yes	303 (59.4)	77 (53.1)		136 (25.2)	15 (25.9)	
No	207 (40.6)	68 (46.9)		403 (74.8)	43 (74.1)	
Treatment arm, n (%)[a]			0.25			0.99
Group A	260 (51.0)	66 (45.5)		269 (49.9)	29 (50.0)	
Group B	250 (49.0)	79 (54.5)		270 (50.1)	29 (50.0)	
Any adverse event, n (%)			0.03			0.12
Yes	314 (61.6)	75 (51.7)		234 (43.4)	19 (32.8)	
No	196 (38.4)	70 (48.3)		305 (56.6)	39 (67.2)	
Surgical retreatment, n (%)[b]			0.004			0.66
Yes	15 (2.9)	12 (8.3)		14 (2.6)	2 (3.4)	
No	495 (97.1)	133 (91.7)		525 (97.4)	56 (96.6)	
Age (years), mean (SD)	52.8 (10.1)	48.7 (10.4)	<0.001	53.5 (10.7)	46.6 (11.5)	<0.001
No. of accidents per day, mean (SD)	3.19 (3.03)	3.27 (2.80)	0.78	3.38 (3.04)	2.79 (2.26)	0.07
Total MESA score, mean (SD)	25.8 (7.34)	25.7 (7.63)	0.82	25.7 (7.61)	24.7 (6.97)	0.36
Total IIQ score, mean (SD)	169.1 (100.1)	179.8 (105.2)	0.26	149.6 (97.2)	169.4 (97.8)	0.14
BMI (kg/m^2), mean (SD)	29.7 (5.87)	31.1 (6.90)	0.03	30.2 (6.56)	31.1 (8.17)	0.44

[a] SISTEr study: *group A* pubovaginal sling, *group B* Burch colposuspension. TOMUS study: *group A* retropubic midurethral sling, *group B* transobturator midurethral sling.

[b] Surgical retreatment included bulking agent injection, fascial or synthetic sling performed up to 24 months after study surgery in the SISTEr study and up to 12 months after study surgery in the TOMUS study.

[c] Chi-squared or Fisher's exact tests were used for categorical variables and a two-sample t test for continuous variables.

may have been differences in eligibility criteria and morbidity of interventions. The SISTEr study, which allowed concomitant abdominal surgery, was associated with a higher overall rate of adverse events than the subsequent TOMUS study which did not allow concomitant abdominal surgery and used minimally invasive surgical techniques.

The use of composite outcome variables has advantages and disadvantages. Our findings that certain components of our composite outcome were more likely to be missing may inform other incontinence researchers regarding their choice of these variables individually or within composite outcomes. Questionnaire data and physical finding data (urinary stress test) that could be assessed during a single visit were least prone to missing data, whereas data on variables that required subject effort while away from the research team (pad test, voiding diary) were more likely to be missing. Our finding that missing data was more likely to be associated with younger subjects adds weight to the need to include older subjects

in clinical research. Efforts to promote retention of younger subjects, while improving study adherence, may improve data completeness in this group. Older subjects, in addition to broadening the generalizability of research, appear to contribute complete data more reliably. This supports the national effort to avoid ageism in research, especially research that might be of relevance to the elderly.

Despite the best efforts of the clinical research teams, missing data is likely to occur, requiring robust methods of analysis to assess whether the missing data are random. Our findings suggest that there may be demographic associations that require consideration to avoid unsound conclusions, including incorrect claims of treatment superiority that may be due to differential drop-out. The National Academy of Sciences encourages clinical trial investigators to consider proactive measures to reduce the amount of missing data in all subjects, especially in groups associated with higher rates of missing data [1]. In addition to an on-going assessment of data quality, assessment of patterns of missing data may permit scientifically sound interventions that minimize missing data through the remainder of the trial.

The analytic approach to missing data is not uniform. Incomplete data can challenge the interpretation of trial outcomes, regardless of the analytic method used for missing data. Although we did not compare the various analytic methods used in these studies, the biostatistical literature is filled with cautionary guidance reminding investigators to assess whether missing data is occurring randomly or in patterns that may pose interpretation bias, limiting the impact of the trial itself.

Our analysis had several minor limitations, including a lack of patient perspective regarding the reason for missing data, which could have provided additional information on subject burden, motivation and study design. In addition, we were unable to compare the effect of the different primary outcome assessment time-points in an identically designed trial, other surgical trials or other trials assessing other interventions, such as drugs, behavioral interventions or other incontinence treatments.

The findings of our analysis are strengthened by the high quality of the study designs, multisite participation, sufficient subject retention and quality control for data management by an experienced coordinating center. These results may be used to assist in deciding on the types of data and how they are collected in future trials regarding urinary incontinence.

Conflict of interest Dr. Brubaker received grant funding from the NIH and is a section editor of Up-To-Date. A. Stoddard was a shareholder of Bristol-Myers Squibbs and Johnson & Johnson with no compensation.

H.J. Litman, H. Kim, P. Zimmern, K. Dyer, J. Kusek and H. Richter have no disclosures to report.

Author contributions L. Brubaker: protocol/project development, data collection, manuscript writing/editing

H.J. Litman: management of data analysis, manuscript writing/editing

H. Kim: management of data analysis, manuscript writing/editing

P. Zimmern: protocol/project development, data collection, manuscript writing/editing

K. Dyer: data collection, manuscript writing/editing

J. Kusek: protocol/project development, manuscript writing/editing

H. Richter: protocol/project development, data collection, management of data analysis, manuscript writing/editing

A. Stoddard: protocol/project development, management of data analysis, manuscript writing/editing

Financial support Supported by cooperative agreements from the National Institute of Diabetes and Digestive and Kidney Diseases: U01 DK58225, U01 DK58229, U01 DK58234, U01 DK58231, U01 DK60379, U01 DK60380, U01 DK60393, U01 DK60395, U01 DK60397, and U01 DK60401. Also supported by the National Institute of Child Health and Human Development and Office of Research in Women's Health.

References

1. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT et al (2012) The prevention and treatment of missing data in clinical trials. N Engl J Med 367:1355–1360
2. O'Neill RT, Temple R (2012) The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. Clin Pharmacol Ther 91:550–554
3. Tennstedt S; Urinary Incontinence Treatment Network (2005) Design of the Stress Incontinence Surgical Treatment Efficacy Trial (SISTEr). Urology 6:1213–1217
4. Albo ME (2008) The Trial of Mid-Urethral Slings (TOMUS): design and methodology. J Appl Res 8:1–13
5. Albo ME, Richter H, Brubaker L, Norton P, Kraus SR, Zimmern PE et al (2007) Burch colposuspension versus fascial sling to reduce urinary stress incontinence. N Engl J Med 356:2143–2155
6. Richter HE, Albo ME, Zyczynski HM, Kenton K, Norton PA, Sirls LT et al (2010) Retropubic versus transobturator midurethral slings for stress incontinence. N Engl J Med 362:2066–2076
7. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA (1994) Health-related quality of life measures for women with urinary incontinence: the incontinence impact questionnaire and the urogenital distress inventory. Qual Life Res 3:291–306
8. Herzog AR, Diokno AC, Fultz NH (1989) Urinary incontinence: medical and psychosocial aspects. Annu Rev Gerontol Geriatr 9:74–119